PROSKAUER ROSE LLP
ELISE M. BLOOM (admitted *pro hac vice*)
ebloom@proskauer.com
NEIL H. ABRAMSON (admitted *pro hac vice*)
nabramson@proskauer.com
ADAM M. LUPION (admitted *pro hac vice*)
alupion@proskauer.com
RACHEL S. PHILION (admitted *pro hac vice*)
rphilion@proskauer.com
NOA M. BADDISH (admitted *pro hac vice*)
nbaddish@proskauer.com
JOSHUA S. FOX (admitted *pro hac vice*)
jfox@proskauer.com
Eleven Times Square
New York, NY 10036
Telephone: (212) 969-3000
Facsimile: (212) 969-2900

PROSKAUER ROSE LLP
PHILIPPE A. LEBEL (SBN 274032)
plebel@proskauer.com
2029 Century Park East, 24th Floor
Los Angeles, CA 90067-3010
Telephone: (310) 557-2900
Facsimile: (310) 557-2193

PROSKAUER ROSE LLP
MARK W. BATTEN (admitted *pro hac vice*)
mbatten@proskauer.com
SAMANTHA R. MANELIN (admitted *pro hac vice*)
smanelin@proskauer.com
One International Place
Boston, MA 02110-2600
Telephone: (617) 526-9600
Facsimile: (617) 526-9899

UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA

AARON SENNE, *et al.*,

Plaintiffs,

vs.

OFFICE OF THE COMMISSIONER OF BASEBALL, an unincorporated association doing business as MAJOR LEAGUE BASEBALL, *et al.*

Defendants.

Case No. CV 14-00608 JCS (consolidated with 3:14-cv-03289-JCS)

Hon. Joseph C. Spero

DEFENDANTS' NOTICE OF MOTION AND MOTION TO EXCLUDE PLAINTIFFS' EXPERT DECLARATIONS AND TESTIMONY OF J. MICHAEL DENNIS, PH.D

Date: February 11, 2022
Time: 9:30 am
Place: Courtroom G, 15th Floor

NOTICE OF MOTION AND MOTION TO EXCLUDE PLAINTIFFS' EXPERT DECLARATIONS AND TESTIMONY OF J. MICHAEL DENNIS, PH.D

PLEASE TAKE NOTICE that on February 21, 2022, at 9:30 a.m. or as soon thereafter as counsel may be heard, Defendants[1] will and hereby do move this Court for an entry of an order excluding the survey, expert declarations and any testimony by J. Michael Dennis, Ph.D.

This motion is based on this Notice, the Memorandum of Points and Authorities, the Declarations of Eugene P. Ericksen and Elise M. Bloom, filed concurrently herewith, accompanying Exhibits, the pleadings and records on file with this Court, all matters of which the Court must or may take judicial notice and such evidence and argument as may be presented at or before the hearing on this matter.

[1] "Defendants" refers to all Defendants named in the Second Consolidated Amended Complaint, with the exception of the MLB Clubs that have been dismissed for lack of personal jurisdiction.

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

FEDERAL CASES

MEMORANDUM OF POINTS AND AUTHORITIES

PRELIMINARY STATEMENT

In holding the Main Survey sufficient for class certification purposes in 2017 under a "tailored" *Daubert* analysis, the Court specifically noted that the survey would be the subject of legitimate challenge at a later stage. The Ninth Circuit agreed. That time has come. Now, at summary judgment, *Daubert*'s gatekeeping requirements apply with full force. Dr. J. Michael Dennis's survey and testimony cannot survive that scrutiny. They must be excluded.

On behalf of the Rule 23 classes and FLSA collective, Plaintiffs seek minimum wage and overtime for all hours worked on team-related activities. Indeed, Plaintiffs' narrowing of their class and collective claims—from seeking compensation for all activities to only those team-related activities—is precisely what allowed them to cross the Rule 23 barrier. Now, at summary judgment, Plaintiffs proffer the survey as the exclusive source of all hours worked during spring training, extended spring training and the instructional league, and all hours worked during the California League Championship Season exclusive of game time and travel time. The survey Dennis conducted cannot possibly carry the enormous weight Plaintiffs ask it to. New evidence adduced during recently-concluded expert discovery—including Dennis' deposition, which did not take place until September 2021—confirms that the survey is inherently unreliable in multiple respects and cannot be used at summary judgment or trial as evidence of hours worked, much less hours worked on team-related activities that is the subject of the claims at issue here. It ignored multiple foundational principles of survey design and administration. It asked the wrong questions of too few players, who gave such widely varying answers to the confusing survey that no reliable conclusion can be drawn from it. Implicitly recognizing the incompetence of the data to answer relevant questions, Dennis and Plaintiffs' damages expert, Dr. Brian Kriegler, propose to discard the vast majority of the survey responses, hoping that in cherry-picking the smallest numbers the Court will excuse the flaws that invalidate their results.

First, though Plaintiffs describe their damages as time spent on "team-related activities," the Main Survey does not measure that time. It asks not a single question, in fact, about players' activities at the ballpark. Finding that relevant question too difficult, the survey simply assumes that

any time spent at a ballpark—whether playing a game or watching TV—was compensable work time. Even the general measure of time spent at the ballpark is too difficult to measure, however. To do so, the survey asked players to think back over months or years, tote up daily experiences, and report the time (to the nearest hour) that they most often arrived at, and departed from, the ballpark in spring training, extended spring training, the fall instructional leagues, and the Championship Season. Even assuming players could perform that exercise accurately, the survey simply measured the wrong thing.

Second, players' responses to the survey varied widely, not only from one Club to the next, but even among players on the same Club in the same year. Whether those inconsistent responses arose because players' memories were different, or because they were free to make individual choices about when to come and go, or because different teams had different requirements, the reasons for the variation were simply not of interest to Dennis; he did not even ask respondents to identify the affiliate they were playing for at the time of response. Whatever the reason for the differences in survey responses, those inconsistencies make clear that they are not reliable evidence of "hours worked," as Dennis represents them to be. He testified in deposition that accounting for the differences in responses was simply not part of his job, and he made no effort to do so; instead, he simply aggregated the responses for players on thirty different Clubs and for multiple affiliates within those Clubs into an undifferentiated whole without making any attempt to understand whether the variability of the data permitted him to lump all of the responses together as he did.[2] It has now become clear that we can determine the identity of each respondent, which facilitates an assessment of the variability by and within teams, and demonstrates the fundamental unreliability of the survey.

Third, that deficiency was especially pronounced with respect to the California League. Indeed, Dennis admitted that the survey was not designed to target the California League population and that he had no empirical data that allowed him to reliably extrapolate the survey responses he did receive to players in the California League. As it turns out, fewer than a dozen respondents in the sample Dennis and Plaintiff's damages expert, Dr. Brian Kriegler, ultimately relied on had

[2] In fact, Dennis conceded that he could not have conducted a quantitative assessment of the varying responses even if he had wanted to, because there were not enough survey responses from each team to permit such an analysis.

participated in the California League in 2015, the only Championship Season which this smaller sample was asked about. Worse, all but one or two of those players had also played in other leagues during the same season, and the survey made no effort to determine which league the respondents were thinking about when they answered the questions. When asked how he could be sure that the wide discrepancies in responses permitted a conclusion that the minuscule sample of California League responses was representative of the larger population, Dennis had no answer based on statistical analysis; he had not performed any. He pointed instead to his "experience" – even though this is the first wage and hour case of his career – and his "understanding" that minor league players' days were all uniform and routine – even though his own survey data said the opposite.

Fourth, the Main Survey suffers from all of the same fundamental methodological flaws that led the Court to exclude the Pilot Survey in July 2016, and new evidence developed since then confirms the depth of these biases. The Main Survey, like its predecessor, bases its conclusions on a sample that significantly underrepresents foreign-born players, non-opt-ins, and other subgroups of the population, skewing the results. It strains respondents' ability to recall minute, insignificant details of old events. Though it asked 65 burdensome, confusing questions, it told respondents they could complete the survey in 15 minutes. And it attempted to obtain neutral factual information from respondents who have a direct financial interest in the survey results, some of whom contacted Dennis to ask whether the survey was connected with the litigation, as they suspected.

RELEVANT BACKGROUND

Dennis originally conducted a Pilot Survey in an effort to measure time spent by minor league players on "baseball-related activities." (*See* Dkt. 498.) By its July 21, 2016 Order, the Court granted Defendants' motion to exclude the Pilot Survey and related testimony from Dennis, concluding that the survey was inherently unreliable and employed flawed methodologies. (Dkt. 687, at 96-103.) Citing primarily concerns about self-interest bias and recall bias, the Court held that while some of the problems with the Pilot Survey might theoretically be addressed with changes in a subsequent survey, other problems were "fundamental and cannot be fixed." (*Id.*)

Before the Court's Order was issued, Dennis designed and issued the Main Survey; he collected responses from July 9-27, 2016. (Dkt. 696 ¶ 3.) By the time the Court's Order was issued,

Dennis could not amend the survey to address any of the Court's concerns, and he opted not to change the Main Survey in light of the Court's holding mid-way through its administration, nor to send another survey. (*See* **Exhibit 1**, 84:7-85:6).[3]

Dennis received 720 responses to the Main Survey. (Dkt. 696 ¶ 3.) 93% of survey invitees who had chosen not to opt in to the FLSA collective chose not to participate, as did 75% of opt-ins. (Dkt. 726 ¶ 47.) While more than 40% of the survey invitations were sent to players with foreign addresses, just 10% of the total responses came from those invitees, even though Dennis offered a Spanish-language version of the survey. (**Exhibit 2** ¶ 80.) There were very few responses, typically 25 or fewer, from players on any particular Club, and even fewer for any particular affiliate. (Ex. 2 ¶ 58; **Exhibit 3**, at 8, Appendix B, Tables B1-B5.)

Though Plaintiffs claim it was intended to count time spent on baseball-related activities for purposes of a damages calculation, the Main Survey did not ask any respondent about time spent on such activities. Instead, it asked players when they most often arrived at departed from the ballpark on different types of days, assuming in effect that all time spent at the ballpark was at the behest of management and spent on baseball-related activities. (Ex. 1, 99:25-101:13.) The responses to those questions varied widely from Club to Club, and even among players on the same Club. (*See, e.g.*, Daubert Declaration of Eugene P. Ericksen ("Ericksen Daubert Decl.") ¶¶ 26, 38; Ex. 2 ¶¶ 28-29; Ex. 3, Appendix B.)

By Order dated March 7, 2017, the Court certified three classes, conditionally certified an FLSA collective, and declined to exclude the Main Survey, but noted that the Order "will not preclude Defendants from challenging the sufficiency of the Main Survey and Plaintiffs' damages model on summary judgment and/or at trial." (Dkt. 782 at 56.) On appeal, the Ninth Circuit held that the Court did not abuse its discretion in declining to exclude the Main Survey at the class certification stage, but noted that in doing so the appellate court did not mean "to minimize defendants' criticisms of the Main Survey." (Case No. 17-16245, Dkt. 91-1 at 49.) Indeed, the Ninth Circuit agreed with Defendants that there "are a number of legitimate questions about the

[3] All exhibits referenced herein are annexed to the Declaration of Elise M. Bloom In Support of Defendants' Motion to Exclude Plaintiffs' Expert Declarations and Testimony of J. Michael Dennis, Ph.D. and are referred to as "Exhibit __" or "Ex. __."

persuasiveness of the Main Survey," especially standing alone, but noted that Plaintiffs also intended to rely on team and game schedules, payroll data, and testimony. (*Id.* at 49-50.)

In an implicit acknowledgement of the intractable variability of responses and the deficiencies the Court had noted in the July 2016 Order, Dennis presented results for a "control group," a subset of respondents defined as those who had not opted in to the FLSA collective and had played baseball in 2015 or 2016. (Dkt. 696 ¶ 13.) Those 284 responses became the basis for Kriegler's damages calculations. (**Exhibit 4** ¶ 104.)

After the Court compelled Plaintiffs to produce the identities of survey respondents during expert discovery, on August 27, 2021 (*see* Dkt. 956), that information confirmed that, in the subset of survey respondents on which Plaintiffs' damages model is based, only 9 or 10 respondents (depending on the survey question) participated in the California League in 2015, the only year they were asked about. (Ex. 2, ¶ 60; Ex. 3, at 7-8.) But because most of those players also played on other affiliates that year, it is impossible to tell whether they were reporting on their experience in the California League or some other league. (*Id.*) In fact, only one or two respondents (again, depending on the question) spent the entire year in the California League in 2015 and therefore only those clearly were reporting arrival and departure times in the California League. (*Id.*)

ARGUMENT

A. DAUBERT'S GATEKEEPING TEST IS MORE DEMANDING NOW THAN AT CLASS CERTIFICATION AND PRECLUDES USE OF THE SURVEY ESPECIALLY IN LIGHT OF DENNIS'S TESTIMONY AND OTHER EVIDENCE ADDUCED DURING EXPERT DISCOVERY.

Plaintiffs bear the burden to prove that Dennis's survey is the product of reliable principles and methods, applied to sufficient facts and data, in a way that will help the trier of fact understand the evidence or determine the facts. Fed. R. Evid. 702. The Court acts as a "gatekeeper," ensuring that Dennis's "reasoning or methodology . . . is scientifically valid and . . . [that] that reasoning or methodology properly can be applied to the facts [at] issue." *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 592-93 (1993). The test has two steps: Plaintiffs must show that Dennis's survey is scientifically valid, and that it will assist the trier of fact. Scientific validity, in turn, considers such factors as "1) whether the methodology can be or has been tested; 2) whether the theory and technique has been subjected to peer review; 3) if a 'particular scientific technique' is involved, the

known or potential rate of error; and 4) the degree of acceptance in the relevant scientific community." *Base v. FCA US LLC*, No. 17-CV-01532-JCS, 2019 WL 1117532, at *4 (N.D. Cal. Mar. 11, 2019) (Spero, J.) (citing *Daubert*, 509 U.S. at 592-94).

Because expert testimony "can be both powerful and quite misleading," "a federal judge should exclude scientific expert testimony under the second prong . . . unless he is convinced that it speaks clearly and directly to an issue in dispute in the case." *Id*. at *5 (quoting *Jones v. U.S.*, 933 F. Supp. 894, 900 (N.D. Cal. 1996) (quotations omitted)). This requirement is often referred to as one of "fit" – ensuring that an expert's work is not merely scientifically valid in general but that there is "a valid scientific connection to the pertinent inquiry" such that the testimony will "logically advance[] a material aspect of the proposing party's case." *Id*. (quoting *Daubert*, 509 U.S. at 591; *Daubert v. Merrell Dow Pharms., Inc*., 43 F.3d 1311, 1315 (9th Cir. 1995)).

As this Court recognized at the time, the full force of *Daubert*'s gatekeeping scrutiny did not apply when the Court initially considered Defendants' challenge to Dennis's testimony in 2016 and 2017, in connection with Plaintiffs' motion for class certification. Then, the Court applied only a "tailored *Daubert* analysis which scrutinize[s] the reliability of the expert testimony in light of the criteria for class certification and the current state of the evidence." (Dkt. 687 at 97 (quoting *Rai v. Santa Clara Valley Transp. Auth*., 308 F.R.D. 245, 264 (N.D. Cal. 2015) (citations omitted)).) Moreover, that analysis "is not a final conclusion that will control the admissibility of the expert's testimony at trial." *Cholakyan v. Mercedes-Benz, USA, LLC*, 281 F.R.D. 534, 542 n.53 (C.D. Cal. 2012). Now is the time to apply the full scrutiny that *Daubert* mandates.

At class certification, for example, the Court's consideration of individualized damages issues is limited. (*See, e.g*., Case No. 17-16245, Dkt. 91-1 at 49 ("the need for individual damages calculations does not, alone, defeat class certification") (quoting *Vaquero v. Ashley Furniture Indus., Inc*., 824 F.3d 1150, 1155 (9th Cir. 2016)).) Now, however, as explained in greater detail in Defendants' motion for summary judgment, the question is not merely whether class members have been injured, but by how much: Plaintiffs must show that they have a workable model on which a jury could reliably determine damages for each class member. Because the scope of the inquiry is

now expanded and the legal test more rigorous, the consideration of deficiencies in Dennis's survey that were largely deferred at the class certification stage now require careful scrutiny.

Full development of the evidence over the last several years now demonstrates that the profound limitations of Dennis's survey preclude its use, because it is not scientifically valid, and does not "speak[] clearly and directly to an issue in dispute in the case." *Base*, 2019 WL 1117532, at *5 (quoting *Jones*, 933 F. Supp. at 900). Dennis' testimony and other evidence must be excluded.

B. DENNIS CONFIRMED THAT THE SURVEY DOES NOT MEASURE THE TEAM-RELATED ACTIVITIES THAT ARE THE SUBJECT OF PLAINTIFFS' CLAIMS.

Seeking to remedy deficiencies in their original proposed classes, Plaintiffs argued that their revised claims concern only "team-related" activities. To be relevant and admissible, therefore, the Main Survey must reliably measure time spent on those activities. That is because an essential element of the *Daubert* test is ensuring a "fit" between the expert evidence and the issues to be decided. "This requirement is more stringent than the relevancy requirement of Rule 402 of the Federal Rules of Evidence, 'reflecting the special dangers inherent in scientific expert testimony.'" *Volterra Semiconductor Corp. v. Primarion, Inc*., 2013 WL 6905555, at *15 (N.D. Cal. Nov. 18, 2013) (Spero, J.) (citation omitted).

Dennis's survey, however, did not ask a single question about "team-related activities," or, for that matter, the nature of any activity that a player engaged in while at the ballpark. Dennis testified that Plaintiffs' counsel "asked me to answer . . . how much time did these Minor League Baseball players spend at the ballpark," (Ex. 1, 28:21-23) – the very same topic investigated by the Pilot Survey, which the Court excluded. Instead of asking directly about those activities, however, in the Main Survey Dennis asked only about the times that players most often arrived at, and departed from, the ballpark, without regard to what they did during the intervening hours apart from meals. (Ex. 2 ¶ 13(c); Ex. 3, at 12-13.) Thus, despite the supposed improvement in questioning from Pilot Survey to Main Survey, Dennis still did not measure hours "worked." He agreed in deposition that he "asked not one question about how [survey respondents] actually spent their time." (*Id.* 101:5-13.) The survey accordingly does not measure the time spent on activities for which Plaintiffs seek compensation.

As Dennis pointed out in his deposition, this disconnect between time spent on team-related activities, which is now the basis of Plaintiffs' claims, and the "most often" arrival and departure times the survey measured, arose at least in part because the Plaintiffs revised their proposal to focus on team-related activities *after* Dennis had already designed, sent, and begun to receive responses to the Main Survey, which he had designed under Plaintiffs' original, failed approach. (Ex. 1, 70:14-15) ("I produced these numbers at the time where the litigation was national in scope."). By the time the Court issued its July 21, 2016 Order excluding the Pilot Survey, the Main Survey – which attempted to elicit the amount of time spent at the ballpark, not about team-related activities in particular – was already in the field, and Dennis made no changes to it at that point or thereafter. (*Id.* 84:7-85:18.) Thus, the Main Survey collected no more information about "team-related activities" than had the original, excluded Pilot Survey.

In denying Defendants' original motion to exclude the Main Survey on this ground, the Court acknowledged that Dennis's survey data "may or [may] not be sufficient to establish the ultimate issue of how much actual work was performed by the putative classes," but held it sufficient at the class certification stage because, the Court believed, it would provide information as to "*whether* the class members performed work and will provide estimates of the amounts of time they worked." (Dkt. 782, at 42 (original emphasis).) We consider each of those suggestions in turn.

First, at the class certification stage, the limited question of whether Plaintiffs performed any compensable work at all may have been relevant, because at that preliminary stage some degree of imprecision in damages may be tolerable. (*See, e.g.*, Case No. 17-16245, Dkt. 91-1 at 49 ("the need for individual damages calculations does not, alone, defeat class certification").) But the question now is whether Plaintiffs can present a reliable damages model that will allow a jury to assess damages with confidence that they can ascertain hours actually worked.

The Main Survey does not accomplish that purpose or contribute reliably to an analysis that does. Even under the "continuous workday" rule on which Plaintiffs rely to excuse the analytical gap between the survey results and the issue to be determined, an individual's compensable time does not begin on a workday until he or she performs the first task that is "integral and indispensable" to the job. *See, e.g.*, *Bamonte v. City of Mesa*, 598 F.3d 1217, 1222, 1225-26 (9th

Cir. 2010).[4] Under Plaintiffs' theory of liability, the continuous workday is bookended by the first team-related activity and the last. A player who arrives at the ballpark early to socialize with others or watch television has not begun a compensable workday and is not engaged in team-related activities, and a player who remains at the park after all team-related activities are over for the day is not entitled to continued compensation. The survey thus necessarily overestimates "hours worked" to an unknowable degree. (Ex. 2 ¶¶ 13(c), 70; Dkt. 726 ¶¶ 32-33.) Further, even once a compensable workday begins, breaks that are long enough to allow an individual to use the time for his or her own purposes (that is, to engage in non-team-related activities) are excluded from hours worked. *See, e.g.*, 29 C.F.R. §785.16. Dennis conceded in deposition that if, after arriving at the ballpark and engaging in some training activities, a player then "took a break for a couple of hours" to "play cards or watch TV," his survey would nonetheless count those activities as hours worked, because the arrival time was "when my manager expected me to get to work," and "downtime" in between was not captured or considered by him to be relevant. (Ex. 1, 116:9-118:7.)

Second, the Court posited in its March 17, 2017 Order that despite the lack of "fit" between the things Dennis measured and the things Plaintiffs must prove, the survey results might still turn out to be helpful to the jury by providing "estimates of the amounts of time [players] worked," precisely because the survey evidence could be "combin[ed] with other evidence such as the daily schedules and witness testimony." (Dkt. 782 at 42.)

As it turns out, none of that other evidence supplements or refines the inaccurate estimates of work hours that Dennis and Kriegler derive from the Main Survey. Neither expert considered the team schedules or deposition testimony to be independent evidence of hours worked; to the extent they considered those items at all, they did so only in an attempt to validate the conclusions they

[4] State law does not differ for these purposes. California does not impose any requirement that employers pay for time spent at the workplace when the employer does not require the employee to be there. *See, e.g.*, *Overton v. Walt Disney Co.*, 38 Cal. Rptr. 3d 693 (Cal. App. 2006) (time spent riding employer's shuttle from designated employee parking to worksite not compensable because shuttle use was optional). Arizona's administrative definition of "hours worked," Ariz. Admin. Code R20-5-1202, does not impose such a requirement either, but rather merely follows the federal rule that time employees spend when *required to be on duty* but not actively working is compensable. *See* Industrial Commission of Arizona, "Substantive Policy Statement Regarding Interpretation of 'Hours Worked' for Purposes of the Arizona Minimum Wage Act" (August 16, 2007), https://www.azica.gov/labor-substantive-policy-hours-worked.

drew from the survey responses. The schedules and testimony are not offered as evidence of hours worked. (Ex. 2 ¶ 23.) Dennis acknowledged that apart from its use in validation (which was inadequate, see Part C below), this other evidence had little independent value. He testified that printed schedules "sometimes . . . change for the players," and so mismatches between the schedule and the hours reported on the survey would not be surprising. (Ex. 1, 243:12-244:3.) It has now become clear that the survey is the ***only evidence*** that Plaintiffs rely on to determine total "hours worked" for the training seasons and for nearly all of the championship season. (Ex. 2 ¶¶ 20-24.)

Accordingly, whatever value the survey might have offered at class certification to answer the question of "*whether* the class members performed work," (Dkt. No. 782 at 42) that question is no longer the touchstone of admissibility, and the survey results cannot offer representative evidence of "hours worked" because the survey did not purport to measure time spent on team-related activities. Baked into the Survey's basic design was the assumption that a player's compensable workday commenced whenever he arrived at the ballpark, and extended through the time that he chose to leave. That fundamental lack of fit between the Plaintiffs' claims and the things the survey measured renders it unreliable as a measure of hours worked, and more likely to mislead than to inform the jury. And although this Court noted that other record evidence could be used to supplement the survey, Plaintiffs decided not to use any of that other evidence to calculate hours worked. On this ground alone, the Survey and Dennis's testimony about it must be excluded.

C. DENNIS FAILED TO ACCOUNT FOR WIDESPREAD VARIABILITY IN ARRIVAL AND DEPARTURE TIMES THAT PLAYERS REPORTED.

Fundamental to the admissibility of a survey is proof that the responses from the smaller group are representative of that population. *Sec. Alarm Fin. Enters., L.P. v. Alder Holdings, LLC*, 2017 WL 5248181, at *4-7 (D. Alaska Feb. 21, 2017) (excluding survey where expert "g[ave] no indication of how or why [he] concluded that" survey respondents were "fairly representative of the entire population"); Reference Manual on Sci. Evid. 359, 379 (3d ed. 2011) ("Identification of a survey population must be followed by selection of a sample that accurately represents that population."). That is particularly so here, where Plaintiffs seek not only to draw conclusions about the activities of the classes as a whole, but to require payment of damages to every class member.

To draw that inference, under accepted principles of survey research, the sample size needed depends directly upon, among other things, the degree of variability in the population under study. *See, e.g.*, *Duran v. U.S. Bank Nat'l Ass'n*, 325 P.3d 916, 940 (2014) ("It is impossible to determine an appropriate sample size without first learning about the variability in the population.") As the *Duran* concurrence elaborated:

> In other words, a valid sampling plan must take into account individual variation within the population, and in that sense, consideration of individual issues is "baked into" the plan's design. Litigation over the degree or nature of variability in the population may result in a determination that no valid sampling plan would be practical or efficient, that multiple samples must be used in order to capture heterogeneity within the class, or that a sampling plan is viable only for a certain subset of the class.

Id. at 950 (Liu, J., concurring).

As reviewed in detail in the rebuttal reports of Dr. Denise Martin and Dr. Jonathan Guryan, and Dr. Eugene Ericksen's *Daubert* declaration, the Survey results here show wide variation in arrival and departure times – not just across the survey respondents as a whole, but from one Club to the next, and even, strikingly, among players on the same Club. (*See, e.g.*, Ericksen Daubert Decl. ¶¶ 26, 38; Ex. 2 ¶¶ 28-29; Ex. 3, Appendix B.) That variability cries out for explanation, because it bears directly on liability. It may be that respondents answered the questions so differently from team to team because of "differences in expectations of attendance by team and Club level," (Ex. 2 ¶ 59) and that respondents answered differently even within the same team because they were free to make their own choices about when to arrive and depart or because affiliates at different levels had different expectations. Dennis fails to provide this essential explanation.

1. Dennis Conducted No Analysis of Variability and Offers No Explanation for It.

Dennis neither collected information nor performed any analysis by team, by affiliate, by level (AA or AAA divisions, for example), or in any other respect. (Ex. 1, 168:11-16 ("I was not tasked with creating survey results and estimates for subgroups at the team level or at . . . the level of the affiliate that the player was associated with, and that kind of thing.").) Although Dennis knew the identities of the survey respondents, he chose not to conduct this analysis. He testified that he "did not have the objective of designing the survey to produce subgroup estimates," and agreed that analyzing variability by affiliate "wasn't [his] job." (*Id*. 44:10-12, 134:2-21). Similarly, he did not

compare responses from players on the same team in the same year, and testified in fact that he ***could not have done so*** because the data he collected was insufficient: there were so few responses on a per-team basis that the survey results "would not support a quantitative analysis at the team level." (*Id*. 136:9-11.)

Having not conducted, nor even considered, an examination of the variability in the responses by team or affiliate, then – and in fact now conceding that the data he collected does not even permit such an analysis – how could Dennis be confident that the overall results he reported were representative of (as Dennis articulated it, "projectable to") the classes at issue in this litigation, as *Daubert* requires? He pointed to two things: "my expertise and experience" in having conducted many surveys, and his "understanding" that minor leaguers substantially "followed similar routines in terms of what their work expectations were." (*Id*. 169:2-24.) Neither offers any ground for confidence.

First, with respect to his experience, Dennis conceded that this is the first survey he has ever conducted in a wage and hour case, and his first attempt ever to measure hours worked. (*Id*. 19:25-21:2.) His experience is in the very different context of gauging subjective consumer opinions and performing marketing research. (*Id*. 33:4-34:9.) Even if he had relevant experience, hollow requests for blind trust are no substitute for statistical evidence that the collection of survey responses can be applied either to the population as a whole or to any particular subgroup or subclass. *See, e.g*., *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997) ("nothing . . . requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert").

Second, Dennis fares no better in trying to substitute his "understanding," of uncertain origin, that minor leaguers all follow similar schedules in place of plain evidence from the survey reporting that they do not. The seven Detroit Tigers respondents for spring training away games, to take one example, reported hours ranging from 3.25 to 10.5, while the six Tampa Bay Rays respondents' answers in the same circumstances ranged from 6.5 to 9.5. (Ex. 3, Appendix B, Table B6.) In the fall instructional league, when no game was played, the two New York Yankees respondents reported between 1.25 and 4.5 hours, whereas the two Arizona Diamondbacks players who answered the question reported between 6.5 and 7.25. (*Id*. Table B14.)

Pressed to explain the variability, Dennis testified, "Well, I mean, it could be that some players are given different instructions by their supervisors than others." (Ex. 1, 172:9-11.) "It *could* be." But Dennis did nothing to investigate this pronounced variability and instead brushed all of it aside by combining the results from respondents who played for 30 different Clubs for countless different affiliates in several different years into a single aggregate. Yet it is equally likely, so far as Dennis is able to say, that the variation arose because players made individual choices to arrive at the ballpark for their own personal purposes before anyone required them to do so – to spend time with colleagues or watch television, for example – or to hang around the ballpark after all baseball-related activities had ended for the day.[5] Or it may arise because the survey suffers from self-interest bias, recall bias, non-response bias, and other flaws that skew the results and generate variations based on faulty memory or exaggeration. Dennis did not ask questions about the reasons for the variability in responses; nor did he make any attempt to analyze the data to find out its source.

2. Dennis Failed to Validate The Survey Responses.

Dennis argues in his report that he validated the Main Survey responses by comparing them to team schedules and deposition testimony. (Dkt. 696 ¶¶ 26, 29.) Again, though, the comparison to other evidence failed to account for the variability in the responses.

a. The Schedules Dennis Reviewed Do Not Confirm His Results.

First, Dennis compared his survey results to 85 "documents that make reference to scheduled work," mostly daily itineraries. (*Id.* ¶ 26.) As an initial matter, these were not a random sample of itineraries; they were merely the itineraries that Dennis believed were available at the time he was conducting his analysis, since supplemented with hundreds more that he has never evaluated. There is no basis to conclude that the itineraries reviewed are representative of any larger set.

In any event, the schedules selected did not validate the survey responses. The hours derived from these documents, even on an aggregate basis, did not match the hours that Dennis derived from the survey responses. He nevertheless adjudged them to be close enough (again, without any statistical calculations, just his own subjective judgment). (Dkt. 696 ¶ 27.) Dennis concedes that the

[5] There was substantial evidence in the deposition testimony to suggest that players made such voluntary decisions routinely. (Ex. 2 ¶ 48 n.31, ¶ 70 n.55.)

survey responses reported higher hours than shown on the schedules, but concluded that he could ignore the difference on the hypothesis that the schedules did not report all team-related activities. (*Id*. ¶ 28.) In other words, where the "validating documents" did not match his assumption, he simply concluded that the documents must be incomplete. In the end, then, he used the documents to "validate" the survey only when he deemed them close enough to the survey responses, and ignored them otherwise.

Moreover, despite the variability across respondents on different Clubs, Dennis did not compare survey responses from players on a particular team against team schedules for that team in that year – presumably a more reliable way to validate the survey responses. Asked about that possibility in deposition, Dennis agreed that "[i]t sounds like . . . an approach that may have some merit," but acknowledged that he "did not consider that." (Ex. 1, 242:21-22; *see also id*. 43:18-22 ("[I]'m agreeing with your overall statement that it's useful to have information on the different groups within the survey in order to make some expert judgments about the generalizability of the information.").) But he went on to say that he could not have addressed the variability in the data by performing a validation exercise at the team level with the 85 schedules that he selected to look at: "Even if I wanted to do that, with only 85 schedules, that's going to be a limited exercise. How do I even know that those 85 schedules are a representative sample of the schedules?" (Ex. 1, 244:21-25.) That is a good question. And it is a question that Dennis failed to ask or even, apparently, to think about before his deposition. Yet he chose to rely on those 85 possibly unrepresentative schedules to validate his survey work, while at the same time dismissing them for failure to capture all team-related activities. (Dkt. 696 ¶ 28.)

Moreover, Dennis did not even review any Championship Season schedules in conducting his validation exercise, much less one for a team in the California League, which is the only Championship Season class at issue in this case. (Dkt. 696-3.) Dennis offers no basis to conclude that the Championship Season experiences of players in the California League can be validated by itineraries for different portions of the year.

b. The Deposition Summaries Dennis Requisitioned From Kriegler's Team Offer No Validation of the Survey.

Second, Dennis claims in his report to have validated the survey responses by comparing them to deposition testimony. (Dkt. 696 ¶ 29.) Again, the testimony did not match the survey responses; again the survey responses reported higher hours; and again, Dennis chose to disregard the discrepancy, suggesting that perhaps the solution would be to look only at the responses at the tenth percentile and ignore the remainder (as, not coincidentally, Kriegler eventually did). (*Id.*)

Further, Dennis did not select the testimony that he relied on for this purpose, saying that he was "not so lucky" as to have a team to help him with such tasks. (Ex. 1, 238:8-9.) Instead, he "describ[ed] what [h]e needed in the way of data" to Kriegler, and someone on Kriegler's staff ("I don't recall who") picked out ten deposition transcripts and prepared one-sentence summaries of certain testimony, which Dennis then used for his "validation." (*Id*. 239:16-23.) The exhibit that Kriegler's staff prepared, (Dkt. 696-4), quotes no testimony but merely lists a curt description purporting to summarize selected testimony from cherry-picked depositions. The document was titled "Number of Hours that Players Arrive Prior to the Start of Championship Season Games According to MLB Clubs' 30(b)(6) Witnesses," even though none of the witnesses was a Rule 30(b)(6) witness testifying on behalf of a Club, but was merely testifying from personal experience. Dennis could not account for that description because he had not prepared the document. (Ex. 1, 63:17-64:17.) He also did not ask to see the depositions of the named Plaintiffs – explaining that he did not because, confoundingly, "I'd rather do my own research" – nor did he ever ask for a full list of the available depositions that he might review to validate the survey. (*Id*. at 239:24-241:2.)

c. Dennis's Cognitive Interviews Should Not Be Considered, And In Any Case Do Not Validate The Survey.

Finally, Dennis contends that he confirmed the quality of the data generally by conducting eight telephone "cognitive interviews" with survey participants. (Dkt. 696 ¶¶ 43-44.) He conceded in deposition, however, that he has no notes or other documentation concerning these interviews; indeed, his practice is never to take notes. (Ex. 1, 97:17-25 ("A. No. No, I don't take notes for cognitive interviews. Q. So you have no record of that whatsoever? A. No. I relied on my expertise in conducting these cognitive interviews . . . .").) As a result, Defendants cannot assess the

reliability of the interviews – whether Dennis asked leading questions, whether his description of the interviews in his report is accurate, what answers the interviewees gave, whether they made statements inconsistent with Dennis's expectations, and so forth. (*See* Ericksen Daubert Decl. ¶ 42.) His failure to create any record of the interviews whatsoever should preclude their use. (*Id.* ¶ 43.)

In any case, accepted principles of survey research dictate that cognitive interviews be used only to assist in the wording of questionnaires, not to confirm the quality of data. (Dkt. 726 ¶ 51.) Because these interviews were not "double blind" – Dennis knew what conclusions he wanted to reach before conducting the interviews – and because of "expectancy effects" – the tendency for a researcher's prior opinions to influence data collection – the interview results, even if they could be verified, offer no meaningful confirmation of the Main Survey or its results. (Dkt. 726 ¶ 17, 51-52.)

d. Dennis's Failure to Conduct a Pre-Test Further Demonstrates That the Survey Results Cannot Be Validated.

Accepted industry practice also requires a pre-test, in which a small group of respondents takes the questionnaire in the presence of an interviewer, who discusses with them the questions that may be difficult to answer. (Ericksen Daubert Decl. ¶¶ 33-35.) Shari Diamond, in the *Reference Guide on Scientific Methods*, comments that there is a consensus of survey research professionals that pre-testing must be done, and the American Association of Public Opinion Research describes pre-testing as "the only way of finding out if everything 'works.'" (*Id.* ¶¶ 34-35.)

Dennis testified that he performed a pre-test of the Pilot Survey, but admitted that he did not do so for the Main Survey. (Ex. 1, 220:21-221:12.) And while he characterized the Pilot Survey itself as a pre-test of the Main Survey, he also conceded that he changed some questions from one survey to the next, and those changed questions were not pre-tested. (*See id*. 222:5-10.) In fact, the questions new to the Main Survey included those at the heart of Dennis's conclusions and Kriegler's damages calculations: the questions about when respondents most often arrived at and departed from the ballpark for spring training, extended spring training, and the fall instructional league. (Dkt. 696 ¶ 37.) Thus, the key questions used by Plaintiffs' experts in developing "Hours Worked," and therefore the damages calculations on which Plaintiffs rely, were never pre-tested. The failure to do so and the resulting confusion, as described below in Part D.3, render the Main Survey unreliable. As the Court held in *MacDougall v. Am. Honda Motor Co., Inc.*:

> Boedeker failed to conduct a valid pretest survey using the same instrument utilized in the final survey. Instead, for reasons Boedeker does not adequately explain, the final survey given to respondents fundamentally differed from the pretest. . . . This design choice undermines the reliability of Boedeker's report in three fundamental ways. First, and most obviously, the purpose of a pretest survey is to test survey questions and choices and determine whether they are misleading, confusing, or suggestive. The value of a pretest survey is lost, however, when the pretest survey and final survey are substantially different.

2020 WL 5583534, at *7 (C.D. Cal. Sept. 11, 2020); *see also United States v. Dentsply Int'l, Inc.*, 277 F. Supp. 2d 387, 438 (D. Del. 2003), *rev'd on other grounds*, 399 F.3d 181 (3d Cir. 2005) (holding lack of pre-test for expert's survey "renders the results unreliable").

In sum, Dennis's purported "validation" of the survey results did no such thing. His minimal exercise failed to explain or even engage with the variations in survey responses, and where the "validating" documents were inconsistent with the survey results, he chose to ignore them. As *Daubert* holds, "[o]rdinarily, a key question to be answered in determining whether a theory or technique is scientific knowledge that will assist the trier of fact will be whether it can be (and has been) tested." 509 U.S. at 593. The survey is unreliable and must be excluded.

D. THE SURVEY IS ESPECIALLY MEANINGLESS AS TO THE CALIFORNIA LEAGUE.

The survey's shortcomings in ensuring that the responses are representative of the classes as a whole, in accordance with accepted statistical practice, are particularly acute with respect to the California class, the only class that involves player time during the Championship Season. Dennis essentially admits that the survey was not designed to be representative of the California League. In part because Plaintiffs did not propose the California League class until after the Main Survey was under way, Dennis did not design the survey to ensure any particular level of participation by participants in the California League. (Ex. 1, 69:21-70:2; *see id.* 72:6-8 ("[At the time] I wrote this report . . . this litigation development regarding the California League had not occurred.").) Asked how he would have designed the sample if he "had been asked to provide estimates for players in the California League," Dennis first said "clearly I have not thought about that task," but then said that the standard approach – "every sort of researcher will tell you the same thing" – would be to "oversample" California League participants. (*Id*. 159:11-23.) "Typical practice," he went on to say, "is to select those respondents, potentially with certainty," meaning that instead of random sampling, "it would simply be a census of all the California League players that would be eligible.

They would be class members, basically, that are part of the California League lawsuit." (*Id.* 160:22-161:6.)

With no California class to design for at the time, however, that is not what Dennis did. In fact, we now know that of the 284 individuals who make up what Dennis calls the "control group," (Dkt. 696 ¶¶ 5-6, 13-14) which became the basis for Kriegler's damages calculations, just nine or ten participants in the California League (depending on the question) responded to the survey for the single year (2015) the survey asked about. (Ex. 2 ¶ 60; Ex. 3, at 7-8.) Of those, eight also played for other affiliates, outside the California League, during the year surveyed, (*id.*), so it is impossible to know whether those respondents were answering questions about arrival and departure times with their California League play in mind or their time playing elsewhere.[6] Given the rampant and unexplored variability in responses that runs throughout the survey data, nine or ten responses from a single year simply provides no basis for a reliable conclusion about "hours worked" during the California League Championship Season. Pressed on that point, Dennis again admits that there is no empirical data that would allow him to extrapolate this small number of responses to a broader group of California League participants, and instead fell back to a general appeal to his experience, none of which, he concedes, was in the wage-and-hour context:

> Q. But what is the basis for using the national [level] information to calculate damages for the California League? . . .
>
> A. . . . So could there be some differences in start times for games? Could there be some differences in the California League with reference to the start of night games? Those are all possibilities, but my understanding is that these leagues across the country have a lot more in common than they have differences, and that gives me confidence that the national-level data can project to a specific league, like the California League.
>
> Q. But you don't have any empirical data to back that up, do you?
>
> A. I have, you know, 20 years of experience of data in doing consumer surveys . . .
>
> Q. Okay, but you have no empirical data in this case that would support that; is that right?

[6] In his deposition, Dennis was asked, "But if [respondents] were part of two [clubs], how do they know which team to answer on behalf of?" Dennis responded, "Yeah, it's a good question." He went on to suppose that the respondent would select "the team that they played for the most that year," but acknowledged that was "an assumption that I'm making there." (Ex. 1, 194:22-195:9.)

A. Well, I have the empirical data from all the work I've done over the course of my career on a wide range of topics. That's what I have.

Q. Right, but nothing specific to the wage and hour context; correct?

A. No, I established earlier that I do not have that in my portfolio. . . .

(Ex. 1, 182:5-184:11.) Thus, Dennis's conclusion that the survey results may be applied reliably to the California class is pure *ipse dixit*. In the end, Dennis conceded that the survey did not measure hours worked by participants in the California League: "Q. . . . there's no way to measure hours worked for the California League using your survey; correct? A. I – I think what I said earlier, it's something that I have not looked into yet. . . ." (*Id.* 184:12-21 (objection omitted).)

Dennis's inadequate validation of the survey as a whole was also particularly deficient with respect to the California League. None of the itineraries that he used in purported validation, as described above in Part C.2.a, were from the California League; indeed, Dennis reviewed ***no Championship Season itineraries at all***. (Dkt. 696-3.) The only California-specific evidence he used was some of the deposition excerpts that Kriegler's staff picked for him, and Dennis testified he did not know whether those excerpts related at all to the California League. (Ex. 1, 69:16-70:2.)

With very few respondents and virtually no validating evidence from the California League (or even from the Championship Season generally), the unexplained variability in the data precludes any reliance on the survey results to draw any meaningful conclusions about "hours worked" for the California class. In *Security Alarm Fin. Enterprises, L.P.*, the court excluded a survey for failure to demonstrate why the expert "concluded that these 19 individuals were fairly representative of the entire population of 661" and instead "simply referenced his experience in the field, without analysis." 2017 WL 5248181, at *4. The situation here is even worse, of course, because Dennis seeks to extrapolate from a national sample that includes only nine or ten California League respondents, only one or two of whom were definitely answering with respect to their time in the California League, and apply that sample to a class of over 2,000 players. He does so almost entirely based on survey respondents who never participated in the California League, thereby "includ[ing] in his results, and dr[awing] his conclusions based on, respondents who were not shown and cannot be shown to belong to the target universe." *Id.*, at *7; *see also Kwan Software Eng'g,*

Inc. v. Foray Tech, LLC, 2014 WL 572290, at *4 (N.D. Cal. 2014) (excluding survey for failure to demonstrate "that the survey examined the proper universe of consumers").

The tiny number of California League respondents also had to confront a deeply ambiguous question, central to Dennis's and Kriegler's analysis of the Championship Season, and therefore to the damages calculation for the California class. As to Home and Away Games, the Main Survey asked respondents about their most frequent arrival time "at their stadium or the opposing team's stadium," without any guidance about which the respondent should use. If the player reported to the home stadium and then traveled by bus to the opposing team's stadium, which arrival time should the player report? In deposition, Dennis testified that he assumed respondents would "read the survey question and process it and put it in their context," answering the question based on "their understanding of when their workday is starting." (Ex. 1, 114:7-115:15.) Dennis took no steps, however, to ensure that respondents would all have the same understanding and answer the question the same way, and conceded that he made no effort to test the clarity of the question by checking for variation in arrival times for Home Games as compared to Away Games. (*Id.* 103:17-23.)

Not surprisingly, then, the data shows that answers to these questions varied considerably. (Ex. 2 ¶ 74.) Respondents reported arriving to the stadium in a wide range, from less than two hours to as much as seven hours before the game, and the distribution of those responses was markedly different for Home Games as compared to Away Games. (*See id*.) Martin concludes that some respondents were likely reporting the time that they arrived at their own stadium for Away Games, while others reported the time they arrived at the opposing team's stadium. (*Id*. ¶ 77.) This would not be the first time that a Dennis survey result was excluded for ambiguity in a critical question. *See In re Motor Fuel Temperature Sales Practices Litig.*, 2012 WL 13050523, at *1, *7 (D. Kan. Feb. 29, 2012) (finding a central question ambiguous and therefore that Dennis's survey "does not meet the standards for reliable survey research"). The survey simply cannot be relied upon to describe the experience of the California class, and must be excluded for this purpose.

E. POOR SURVEY DESIGN RESULTED IN BIASES THAT INVALIDATE THE SURVEY.

The record now demonstrates that the Court's concerns about recall bias and self-interest bias (Dkt. 687 at 103) are at least as fundamental as the Court perceived them to be in 2016, and that

other problems separately render the Main Survey inadmissible, particularly under the full *Daubert* standard that now applies for purposes of summary judgment and trial, rather than class certification.

1. The Evidence Now Demonstrates That Non-Response Bias Prevents The Survey From Accurately Representing the Class.

Non-response bias arises when a potential survey respondent's decision about whether to answer the survey is correlated with the information the subject would report. (Dkt. 726 ¶¶ 15, 44-48.) For a sample of responses accurately to represent the population as a whole, the pattern of non-responses must be random.[7] That risk is particularly acute where, as here, the response rate is so low. Though Dennis invited thousands of players to participate, 93% of non-opt-ins, whom Dennis concluded were most likely to provide reliable responses, and 75% of opt-ins refused to participate, resulting in a response rate of less than 10% overall. (Dkt. 726 ¶ 47.). That low return rate suggest problematic non-response bias. *See Autozone*, 2016 WL 4208200, at *17-18.

And while Dennis dismissed the problem of non-response bias by comparing respondents and non-respondents on the basis of age, fielding position, most recent year played, and number of games played (Dkt. 696 ¶ 9), he nowhere explains why he chose these variables (other than the fortuity that he had access to that data) or why that comparison should provide any confidence that non-respondents were random. In fact, the evidence shows troubling patterns of non-response.

Most prominently, the survey responses significantly underrepresent foreign-born players. While 42.2% of the surveys went to foreign addresses, they make up only 10.3% of respondents, even though Dennis made available a Spanish-language version of the survey. (Ex. 2 ¶ 80.) Further, using a foreign address as a proxy would omit any foreign-born player now living in the United States. Dennis concedes that he made no effort to account for this dramatic gap between the percentage of foreign players invited to participate and those who completed the survey, nor gave any thought to the possibility that respondents from other countries might answer the survey questions differently than respondents born in the United States. (Ex. 1, 144:25-145:14.)

[7] *Marlo v. UPS*, 251 F.R.D. 476, 485 (C.D. Cal. 2008) (surveys must "take measures to assure that nonresponses are random and provide analysis of the reasons for nonresponse"). *See also Wallace v. Countrywide Home Loans Inc.*, 2012 WL 11896333, at *4 (C.D. Cal. Aug. 31, 2012); *In re: Autozone, Inc.*, 2016 WL 4208200, at *17-18 (N.D. Cal. Aug. 10, 2016) (same).

Nor, as discussed above in Part D, did Dennis demonstrate that the survey is representative of California League players, a matter of acute concern because only nine or ten respondents in the "control group" had participated at all in the California League. Eight of those had also played in other leagues, and the survey does not ensure that even those few respondents were answering arrival and departure time questions with their California League experience in mind. (Ex. 2 ¶ 81.) Dennis conceded that he could not be certain. (*See supra* n.2.)

2. Recall Bias, Inherent in the Main Survey, Renders It Hopelessly Unreliable.

The Main Survey asked respondents to recall the times they most often arrived at and departed from the ballpark, generally months and in some cases years before they answered the survey questions.[8] Survey research literature concludes, however, that respondents cannot accurately recall non-salient events such as these for more than two weeks to one month. (*See* Dkt. 726 ¶¶ 4, 7, 9, 13, 19, 23-24, 27-31.)

Dennis seeks to avoid these accepted principles by arguing that minor league players' activities were based on routines, and he posits that routines are more memorable for a longer period. (Dkt. 696 ¶ 32.) As with his failure to account for the variability in the survey responses, however, Dennis again adopts a theory and tries to fit the survey data to it, rather than allowing the data to dictate the conclusion. The survey responses show that arrival and departure times are, in fact, anything but routine, varying from team to team and even from player to player; and the survey does not measure the degree to which those times may have varied from day to day for a particular player, because it only asked respondents to distill their experience into the times they "most often" arrived and departed. Dennis again refers generally to "deposition testimony from named plaintiffs and schedules produced by the defendants" to conclude that minor leaguers' activities are a matter of routine, (*id.*) even though he testified that "as a practice, I don't rely on named plaintiffs' deposition testimony," and even though he otherwise concluded that the schedules are inaccurate and subject to change. (Ex. 1, 240:5-7, 243:23-244:3.)

[8] In particular, all of the "control group" responses concerning the Championship Season relate to a single year, 2015 – meaning that all of them were answering questions about arrival and departure times that had occurred approximately a year earlier. (Ex. 2 ¶ 67.)

In any case, what Dennis claims to have been "routine" is not the arrival and departure times, but the activities themselves – playing in games, taking batting practice, and so on. (Dkt. 726 ¶¶ 12, 23-26, 33.) Even if those activities were sufficiently routine to be recalled more readily, which Dennis does not establish, that offers no confidence that the respondents also recall accurately the times they arrived at and departed from the ballpark to engage in those activities. (*See id.*)

Dennis also fails to account for the irregularity of players' schedules. Some days there are games, and sometimes not; on game days, the location and times vary; some games are at home and some are away. The literature suggests that where the events inquired about are not only non-salient but also irregular, recall is particularly difficult. (*Id.* ¶ 26.)

Dennis neither accounts for nor even acknowledges these complications; like Kriegler, he simply assumes instead that if he ignores most of his own data and looks only at a narrow subset – just 284 individuals out of 720 responses to a survey sent to 7,806 players to represent a population of more than 20,000 – that will excuse the recall difficulty. But Dennis does not explain why the memories of this subgroup should somehow be more reliable than others', and there is no other reason to expect that they would be. It does not matter that this so-called "control group" had played in 2015 or 2016; the survey was taken from July 9-27, 2016, and by then, the literature indicates, none of the respondents could be expected to recall arrival and departure information for their last Spring Training, Instructional Leagues, or any part of the Championship Season prior to July 9, 2016. (*Id.* ¶¶ 2, 38-39.)

Dennis's use of "aided prompt questions," an addition to the Main Survey that asked respondents where they lived, who they lived with, and other details in hopes that thinking about those things might assist respondents in remembering their arrival and departure times, did not serve to address the recall bias problem. Dennis himself concedes that such questions only have the potential to assist where the questions of interest seek routine information, and there is no basis to draw that conclusion about arrival and departure times. In any case, the literature posits that such "aided recall" questions only extend respondents' memories by a month at most. (Dkt. 726 ¶¶ 13, 27-31.) There is no reason to believe the questions here had any greater effect, and Dennis offers no data or other empirical evidence to suggest that they did. (*See id.*)

3. The Burdensomeness of The Survey Contributed to Recall Difficulties.

In its July 21, 2016 Order, the Court acknowledged Ericksen's concern that the Pilot Survey imposed too great a cognitive burden on respondents, leading them to give a "best guess" or to "satisfice" in order to get through. (Dkt. 687 at 99.) The Court suggested that Dennis "reframe the survey questions" and break them down into "more bite-size questions." (*Id.*) Dennis did neither of these things. (Ex. 1, 84:7-85:7.) Though the questions changed in some ways from the Pilot Survey, the Main Survey was equally burdensome. (Dkt. 726 ¶¶ 5, 12-13, 20-22.) The Main Survey asked 65 questions, many of them complex, but instructed respondents that they should be able to complete the entire survey in 15 minutes. (*See id.* ¶ 4.) As one measure of that burden, respondents tended to skip questions that appeared later in the survey, demonstrating fatigue. For the "control group," which Dennis argues is "best able to produce reliable and conservative survey estimates," (Dkt. 696 ¶ 5), Martin found that the percentage of respondents skipping the Spring Training game day arrival question more than doubled from the first to the fourth or fifth module. (Ex. 2 ¶ 77.) That pattern was especially telling because the survey presented the modules in the sequence from longest ago to most recent – that is, from hardest to remember to easiest. (*See* Dkt. 696-5.) So even though the questions became easier as respondents went along – or, at least, asked about more recent events – more respondents skipped questions later in the survey than in the first module. (Ex. 2 ¶¶ 76-77.)

4. Self-Interest Bias Skewed the Results of the Main Survey.

Dennis sought to solve the problem of self-interest bias by comparing overall responses to a so-called "control group." (Dkt. 696 ¶ 5.) The Court had already held that such a comparison would fail to solve the problem, "because all minor league players who played within the applicable limitations periods will likely have a vested interest in the outcome of this action regardless of whether or not they opted in to the FLSA collective." (Dkt. 687 at 101.) Participants perceived self-interest in the survey in other ways, as well. Named Plaintiff Kyle Johnson testified, for instance, that he believed answering the survey would be helpful to the lawsuit. (**Exhibit 5**, 288:22-289:7.) Players' financial incentive to exaggerate their responses is a concern missing from the consumer surveys that comprise all of Dennis's experience before this case.

With respect to non-opt-ins, Dennis contended, despite the Court's holding, that "non-opt-ins did not have a financial interest in – in any way in terms of the results of my survey." (Ex. 1, 124:4-7.) There is no reason to believe that non-opt-ins have any less financial interest in the outcome of this litigation – or in inflating the numbers generated by the Survey – than do those who opted in, particularly considering that state law remedies may offer the prospect of greater relief than is available under the FLSA. Kriegler's calculations for the California class, for example, include that state's higher minimum wage, daily overtime, double time in certain circumstances, wage statement penalties, and so on, none of which is available under the FLSA. (*Cf.* Ex. 4 at Ex. 3.)

Further, Dennis acknowledged that he had received unsolicited emails from "a handful" of participants who believed that the survey was connected with the litigation. (Ex. 1, 79:15-80:7.) The emails confirm that despite the "neutral study branding," (*id*. 78:24-79:3) respondents at least suspected that the survey was connected with the litigation, and therefore that the survey was asking questions on matters in which they had a financial interest. The survey correspondence Dennis sent helped to give that impression, thanking invitees for "making your voice heard." (**Exhibit 6**.) There also was substantial publicity about this litigation around the time the Main Survey was issued. (*See* **Exhibit 7** (sample of public articles around the time that the Main Survey was conducted).)

Separately from bias arising from interest in the outcome of the litigation, the survey literature recognizes the tendency of respondents to inflate their recollections of the extent to which they engaged in certain behaviors perceived to be desirable, such as "work." (Dkt. 726 ¶ 34.) Dennis made no effort to account for this source of self-interest bias, nor could he have. Alone or in combination with the other structural deficiencies, self-interest bias unavoidably skews the survey results and precludes their use in this case. *Risinger v. SOC LLC*, 2018 WL 4258500 (D. Nev. Sept. 5, 2018).

CONCLUSION

For the foregoing reasons, Defendants respectfully request that the Court grant Defendants' motion and exclude from evidence the survey conducted by J. Michael Dennis, Ph.D., the results of the survey, and any declarations or testimony by Dr. Dennis.

Dated: October 29, 2021

PROSKAUER ROSE LLP
ELISE M. BLOOM (*pro hac vice*)
NEIL H. ABRAMSON (*pro hac vice*)
ADAM M. LUPION (*pro hac vice*)
MARK W. BATTEN (*pro hac vice*)
RACHEL PHILION (*pro hac vice*)
NOA M. BADDISH (*pro hac vice*)
JOSHUA S. FOX (*pro hac vice*)
PHILIPPE A. LEBEL
SAMANTHA R. MANELIN (*pro hac vice*)

By: */s/ Elise M. Bloom*
Elise M. Bloom
Attorneys for Defendants